REMARKS

It is noted that claims 14-27 are allowed. Claims 12 and 13 are cancelled.

Claim 11, as amended, calls for:

" A method of source normalization for modeling of speech comprising the steps of:

providing an initial set of speech recognition models;

providing training data collected in a plurality of environments and

training the set of speech models jointly with a set of transformations parameters using a maximum likelihood optimization method such that the set of transformation parameters capture the effects of environmental variations and the speech recognition model parameters capture the effects of speech variations."

Claims 11-13 are rejected under 35 U.S.C. 102 (b) as being anticipated by Woodland et al. Iterative unsupervised adaptation using maximum likelihood linear regression). Claims 12 and 13 are canceled. Claim 11 is amended herein to more clearly present the patentable subject matter over the reference. We wish to point out that a novelty of the subject invention is that of training speech models such that the resulting model parameters are not influenced by environmental factors, but rather capture the important statistical variations of speech. On the other hand, MLLR of the reference is a method of adapting a set of speech models to a particular environment. These are two definitely separate things. Of course, after we train our model parameters, we can

implement MLLR to adapt our models to a particular environment but that is not what the claim is about.

In Woodlands the speech models are determined and left unchanged, then the transformations are determined. During the determination of the transformations, the speech models are NOT changed anymore. Such procedure without changing the original speech models is suboptimal.

In applicant's teaching, we are training speech models such that the resulting model parameters are not influenced by environmental factors. In applicant's teaching we have a method of training HMM models jointly with sets of transformations, wherein the sets of transformations are trained to model the environmental factors, such that the resulting HMM statistical parameters are largely independent of environmental factors.

The following outlines the differences in the steps between that of the application and traditional MLLR.

Adaptation of HMM by MLLR:

1. Train a set of baseline HMMs using some large set of training data collected over multiple environments. The baseline HMM parameters contain statistics that include the variation over the training data environments.

2. Gather some adaptation speech from the environment of interest.

3. Use a Baum-Welch type of processing and the baseline HMMs to determine the likelihood of baseline HMM states at each frame of the adaptation data. (Note that not all baseline HMMs may be used, since the adaptation data may be sparse and may not include speech corresponding to some baseline HMMs.)

4. Use the likelihood mapping in (3) to calculate the parameters **Wi** and **bi** of a set of transformations i, that when applied to the HMMs used in (3) results in maximizing the likelihood increase of the adaptation speech.

5. Apply the transformation set to *all* of the baseline HMMs to form an adapted set of HMMs.

6. Use the adapted HMM model set to perform recognition of other utterances in the same environment as the adapted case.

Applicant's source normalization adaptation:

1. Train a set of baseline HMMs and transformations **Wi'** and **bi'** jointly, yielding a set of HMMs wherein the effect of environment on HMM parameters has been removed by means of joint estimation of the **Wi'** and **bi'**. The resulting baseline HMM parameters contain statistics that are largely independent of the variation found in the training data environments, but rather contain statistics that capture valid acoustical variations in speech.

2. Discard the **Wi'** aand **bi'**, but retain the set of baseline HMMs.

3. Perform steps 2-6 in the traditional MLLR procedure above. The adapted HMMs will model the adaptation environment better, since the baseline HMMs were not contaminated by the variation found in all other environments.

This procedure gives the best possible combination of speech models and transformations. The advantage over Woodlands is illustrated in the experimental results presented in the application.

Since there are no other reasons presented for rejection of the claims applicant's claim 11, as amended, and claims 14-27, these claims are deemed in condition for allowance and an early notice of allowance is respectfully requested.

Respectfully submitted;

Robert L. Troike

Robert L. Troike (Registration No. 24183)

Tel. No. (301) 751- 0825